

Twitter情報分析 空港周辺のトレンド分析

江谷典子

Peach・Aviation株式会社

情報システム開発担当

目次

- 目的
- 開発環境
- 手順
- 分析1 情報量の可視化
- 分析2 話題(トレンド) の可視化
- 分析3 名詞の出現頻度リスト

目的

- SNS Twitter情報分析を行い、Peach号が離陸・着陸する空港周囲の話題（トレンド）を可視化するためにPythonを使ったツールを下記資料を参考にして再構築を行った。

情報処理学デジタルプラクティス報告掲載

「オープンソースによるTwitter検索およびデータ可視化の方法」

https://ipsj.ixsq.nii.ac.jp/ej/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=202959&item_no=1&page_id=13&block_id=8

- その使い方について説明を行い、「観光」「コロナウイルス」を事例とした分析結果を記載する。
- 本ツールはGitHubにて公開する。

<https://github.com/NorikoEtani/PEACH>

開発環境

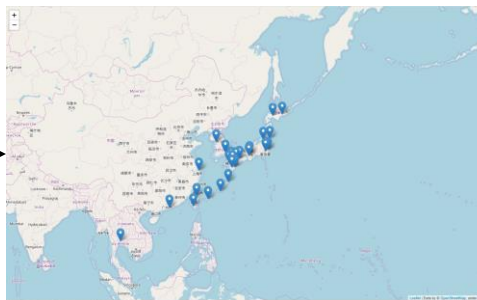
- Windowsのエディション：Windows 10 Home
- システムの種類：64ビットオペレーティングシステム
- Java version 9
- Python 3.5.2 /Anaconda 4.1.1 (64-bit)
- 日本語形態素解析MeCab 0.996

詳細の環境設定は前頁の「オープンソースによるTwitter検索およびデータ可視化の方法」を参照してください。

手順

STEP1

ja_geodata.py



Peach号の利用空港地図
peach_airport.html

PEACH.txt

Peach号の利用空港一覧リスト

geo_peach.txt

空港緯度経度リスト

STEP2 tweetsデータ収集

search_tweet.py

検索キーワード

検索言語（基本設定：ja）

空港からの距離（基本設定：7mi）

各空港別tweetsファイル作成
例）関西国際空港_tweet.txt

STEP3 各空港tweetsデータ可視化

ja_wordcloud.py

各空港別WordCloudファイル作成
例）関西国際空港.png

ja_rank_tweet.py

各空港別名詞出現頻度リストファイル作成
例）関西国際空港_rank.txt

分析 1 情報量の可視化

出力したデータのある場所で「`dir *tweet.txt /o-s`」と入力すると
ファイルサイズ容量の大きい順番に表示されます。

検索キーワード：観光

2020/02/23	03:11	67,926	那覇空港_tweet.txt
2020/02/23	03:11	62,767	熊本空港_tweet.txt
2020/02/23	03:11	61,492	新潟空港_tweet.txt
2020/02/23	03:11	60,767	松山空港_tweet.txt
2020/02/23	03:11	57,273	香港国際空港_tweet.txt
2020/02/23	03:11	54,857	大阪国際空港_tweet.txt
2020/02/23	03:11	53,125	羽田空港_tweet.txt
2020/02/23	03:11	51,325	宮崎空港_tweet.txt
2020/02/23	03:11	49,977	福岡空港_tweet.txt
2020/02/23	03:11	15,348	成田空港_tweet.txt
2020/02/23	03:11	9,874	Suvarnabhumi Airport_tweet.txt
2020/02/23	03:11	7,005	高雄国際空港_tweet.txt
2020/02/23	03:11	5,901	石垣空港_tweet.txt
2020/02/23	03:11	3,958	関西国際空港_tweet.txt
2020/02/23	03:11	3,901	台湾桃園国際空港_tweet.txt
2020/02/23	03:11	3,847	仙台空港_tweet.txt
2020/02/23	03:11	1,582	長崎空港_tweet.txt
2020/02/23	03:11	1,284	大分空港_tweet.txt
2020/02/23	03:11	873	仁川国際空港_tweet.txt
2020/02/23	03:11	503	台南空港_tweet.txt
2020/02/23	03:11	364	金海国際空港_tweet.txt
2020/02/23	03:11	242	新千歳空港_tweet.txt
2020/02/23	03:11	0	上海浦東国際空港_tweet.txt
2020/02/23	03:11	0	奄美空港_tweet.txt
2020/02/23	03:11	0	鹿児島空港_tweet.txt
2020/02/23	03:11	0	釧路空港_tweet.txt

検索キーワード：コロナウィルス

2020/02/23	07:29	70,151	熊本空港_tweet.txt
2020/02/23	07:29	59,957	那覇空港_tweet.txt
2020/02/23	07:29	57,420	香港国際空港_tweet.txt
2020/02/23	07:29	55,400	大阪国際空港_tweet.txt
2020/02/23	07:29	55,393	福岡空港_tweet.txt
2020/02/23	07:29	53,158	新潟空港_tweet.txt
2020/02/23	07:29	51,810	松山空港_tweet.txt
2020/02/23	07:29	38,417	羽田空港_tweet.txt
2020/02/23	07:29	30,996	宮崎空港_tweet.txt
2020/02/23	07:29	13,964	成田空港_tweet.txt
2020/02/23	07:29	12,572	台南空港_tweet.txt
2020/02/23	07:29	4,480	Suvarnabhumi Airport_tweet.txt
2020/02/23	07:29	2,073	仙台空港_tweet.txt
2020/02/23	07:29	1,675	台湾桃園国際空港_tweet.txt
2020/02/23	07:29	1,285	長崎空港_tweet.txt
2020/02/23	07:29	1,170	高雄国際空港_tweet.txt
2020/02/23	07:29	869	石垣空港_tweet.txt
2020/02/23	07:29	0	仁川国際空港_tweet.txt
2020/02/23	07:29	0	金海国際空港_tweet.txt
2020/02/23	07:29	0	釧路空港_tweet.txt
2020/02/23	07:29	0	奄美空港_tweet.txt
2020/02/23	07:29	0	大分空港_tweet.txt
2020/02/23	07:29	0	関西国際空港_tweet.txt
2020/02/23	07:29	0	上海浦東国際空港_tweet.txt
2020/02/23	07:29	0	新千歳空港_tweet.txt
2020/02/23	07:29	0	鹿児島空港_tweet.txt

分析2 話題(トレンド)の可視化

- ①出力したデータのある場所をWindows10のエクスプローラーで表示を行う。
- ②PNGファイルだけの表示を行う。表示を「大アイコン」または「特大アイコン」に設定する。
- ③空港別WordCloudが分かる。

検索キーワード：観光



Suvarnabhumi
Airport



台湾桃園国際空
港



台南空港



金海国際空港



高雄国際空港



仁川国際空港



香港国際空港



熊本空港



大分空港



大阪国際空港



仙台空港



那覇空港



成田空港



長崎空港



松山空港



宮崎空港



関西国際空港



新潟空港



石垣空港



羽田空港



福岡空港



新千歳空港

検索キーワード：コロナウィルス



Suvarnabhumi
Airport



羽田空港



宮崎空港



熊本空港



香港国際空港



高雄国際空港



松山空港



新潟空港



成田空港



石垣空港



仙台空港



台南空港



台湾桃園国際空
港



大阪国際空港



長崎空港



那覇空港



福岡空港

分析 3 名詞の出現頻度リスト

話題の概要がWordCloudで表示されているので、その詳細を知りたい時に名詞の出現頻度リスト「空港名_rank.txt」を参照するように利用してください。